

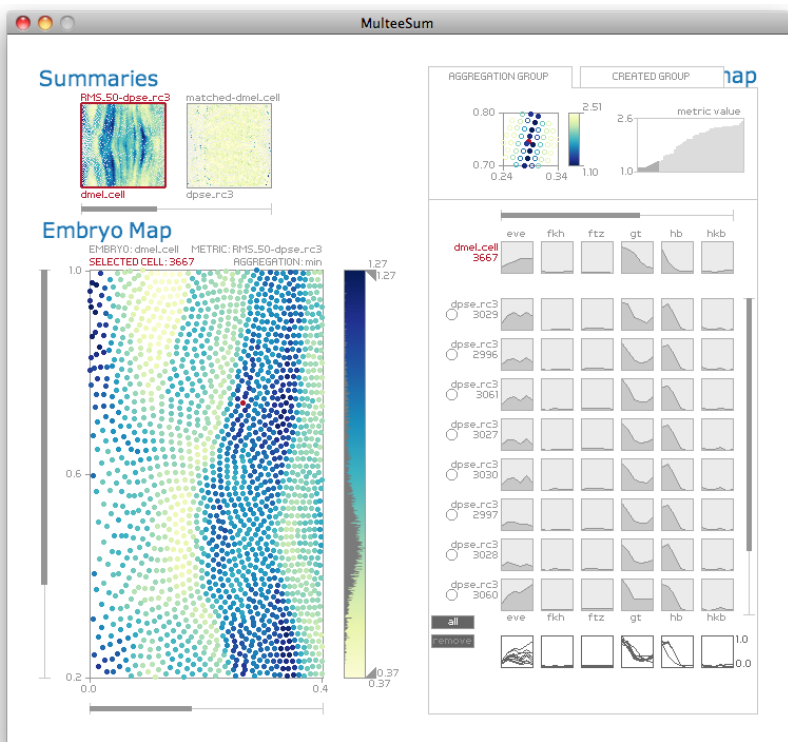
在上周的基础上，继续完成索引部分的实现和时间序列可视化的实现。

1、索引结构

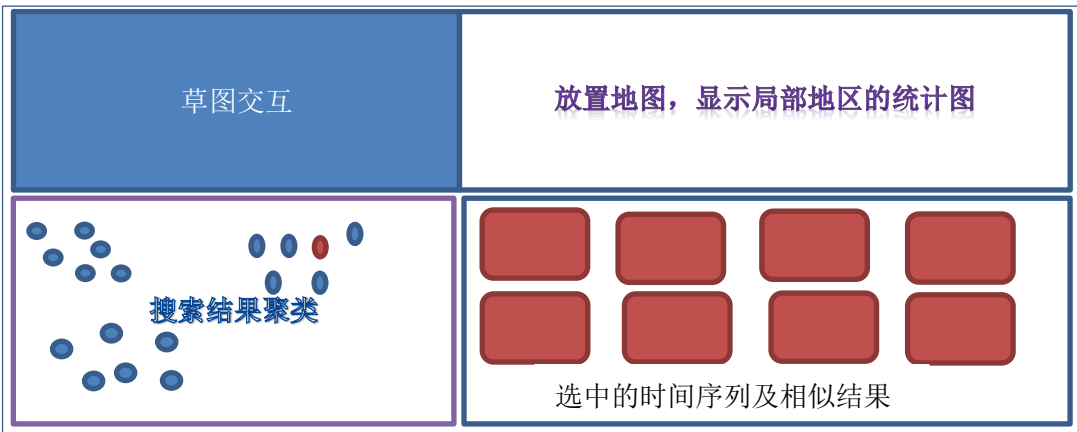
由于 SDL 的索引结构复杂度较高，我思考后决定采用类似八叉树的结构。算法结构比较简单，但是同样也存在一个问题：对于形状处于中间位置的搜索难以实现。也就是说，一个形状，如“stable up down”只是在时间序列的中间出现，而不是以 stable 开头，那么八叉树搜索存在困难。字符串匹配倒是能解决这个问题，但是需要逐个扫描形状字符串，效率也成问题。因此对于 SDL 的索引目前还没有很好的解决办法，仍需要继续思考。简单起见，目前还是在用字符串匹配。

2、时间序列可视化

参考 MulteeSum: A Tool for Comparative Temporal Gene Expression and Spatial Data. 论文中时间序列的可视化实现方法。MulteeSum 采用 Processing 技术，实现方法并不复杂。可视化中有多层次细节，最上层为 Summaries，在 Embryo Map 中可以查看细胞位置，右图可以查看时间序列变化，用户在信息空间中浏览不会迷失。

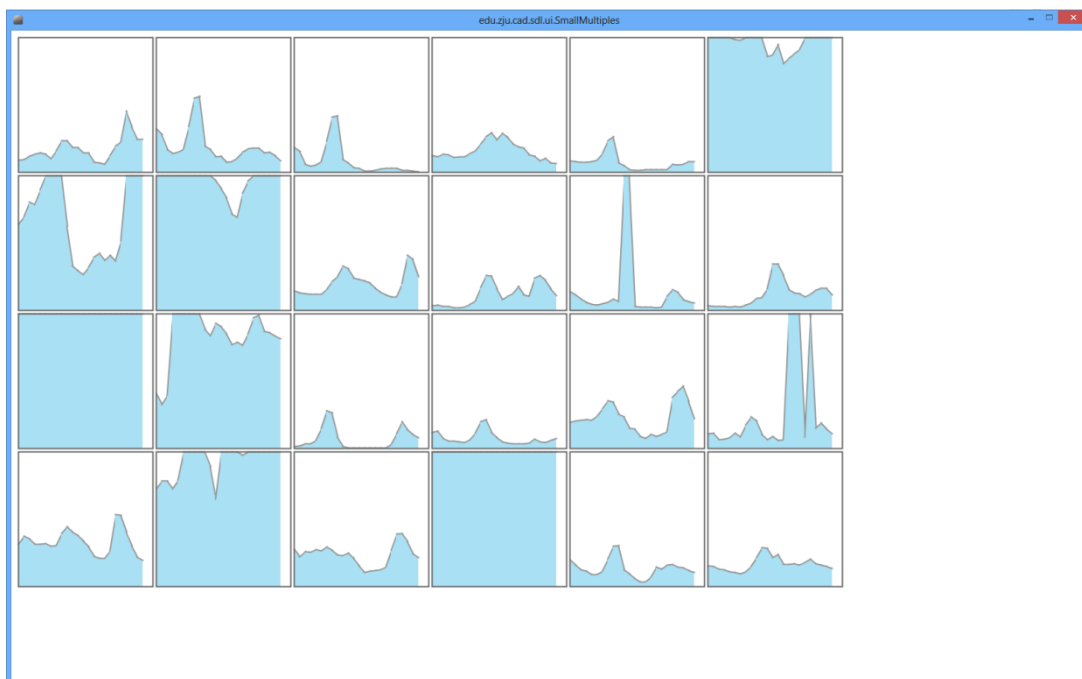


在我们实现的可视化方案中，多层次细节是一个较好的方案。初步考虑可视化方案如下图所示。



- 1) 可以展示搜索结果的总体视图
- 2) 可以展示与某一时间序列相似结果
- 3) 可以展示时间序列在地理空间上的位置
- 4) 搜索结果中的污染物在空间和时间上的统计图

目前已经实现了时间序列的排列显示（未加刻度），搜索“`stable_stable_up_Up_stable_down`”，结果如下图所示。



图中时间曲线有些已经超出了边框，主要原因是数据没有正规化。

3、后续工作

下面要做的是

- 1) 可视化界面的设计和实现
- 2) 聚类的实现
- 3) 寻找恰当的案列，发现数据中存在的模式
- 4) 高效的索引结构